



eCOMMONS

Loyola University Chicago
Loyola eCommons

Computer Science: Faculty Publications and Other
Works

Faculty Publications

11-2006

Electronic Editing of Anglo-Saxon Texts

George K. Thiruvathukal
Loyola University Chicago, gkt@cs.luc.edu

Allen J. Frantzen
afrantz@luc.edu

Recommended Citation

Allen J. Frantzen and George K. Thiruvathukal, Electronic Editing and Anglo-Saxon Texts, Chicago Colloquium on Digital Humanities and Computer Science, 2006.

This Presentation is brought to you for free and open access by the Faculty Publications at Loyola eCommons. It has been accepted for inclusion in Computer Science: Faculty Publications and Other Works by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.



This work is licensed under a [Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 License](https://creativecommons.org/licenses/by-nc-nd/3.0/).
Copyright © 2006 Allen J. Frantzen and George K. Thiruvathukal

Electronic Editing and Anglo-Saxon Texts

Allen J. Frantzen and George K. Thiruvathukal
Loyola University Chicago
Departments of English and Computer Science

Editors of medieval texts that are found in more than one manuscript have traditionally chosen one manuscript, the so-called "best text," and committed it to print, using apparatus such as footnotes to record the readings of other manuscripts where they vary from the manuscript chosen for the printed edition. This expedient, which gave rise to competing theories of textual editing explaining how to weigh the merits of one copy of a text over another, has been rendered superfluous by digital editing. Limitations of space no longer force editors to choose one version over another; editors who wish can transcribe as many manuscript copies of a single text as they wish.

Electronic editing does not, however, solve all problems and creates some problems of its own. Some versions of texts are still better (more complete, more accurate, etc.) than others, and editors need to recognize those differences in evidentiary value. Electronic editing also introduces its own problems. For example, what sort of interface displays multiple versions of a single text most effectively and allows readers to identify the differences among them? How can electronic display allow users who do not read medieval languages to access the material in translation? Digital editions contain an enormous amount of data that can easily overwhelm novice users. It is as important to develop effective navigation tools as it is to provide inclusive data.

Our project approaches a significant corpus of records from the early medieval English church using an innovative approach to display. The texts in question are handbooks of penance written in Old English and used in England between 850 and 1150. Priests consulted these lists of sins and proportional penances when they heard private confession; the handbooks assigned penances for sins ranging from adultery to theft. There are five of these texts in Old English. Each is found in at least three manuscript copies of the late tenth or eleventh century, and one is found in six manuscripts. Only one has been edited in the modern era (i.e., post-World War II) era.

Each version of every text was transcribed and marked up in a version of SGML approved by the Text Encoding Initiative. The markup incorporates such manuscript features as marginal notations, use of color ink for large initials, corrections, and wear and tear on the manuscript page. The display of each page follows what is called a "diplomatic facsimile" format. In this format, the physical appearance of the text is preserved to the extent that typographic conventions allow. Lines of text are broken at the end of the line, words are divided in the edition as they are on the manuscript page, and errors in the text are reproduced without being corrected. Users of the database see a reasonable facsimile of the manuscript page but have numerous options for manipulating the information on the page.

Using this interface, the reader chooses a text. Each sentence of each text has a numerical tag. When the cursor rests on that tag a window automatically opens to provide modern English translating of that sentence and to offer other options. Users can see all possible versions of the sentence in a new window, can see the highlighted sentence and its translation together in a new window, or can switch to the manuscript containing other versions of the highlighted sentence. The process also works in reverse. Users who are consulting the translation of any of the texts can use the highlighted tag from any translated sentence to move to the same options. They can see the original text and translation in a new window or can see all copies of highlighted original text in a new window; in addition, users can switch to any version of the original text. Other supporting data are provided. A cultural index takes users to all sentences concerning, for example, animals, or theft, homicide, and many other topics of interest to a wide range of scholars in the humanities. A glossary offers modern English translations of every Old English word in the corpus. Text and manuscripts are provided with brief explanatory material about their form, content, and historical contexts.

From a technical perspective, there are two distinguishing aspects to this work: *custom transformation*

and *interactivity* (using the ideas of Ajax, Asynchronous JavaScript and XML, but focused on a generated client-side interface). Transformation is ordinarily achieved in XML using the XSLT framework. However, for more complex transformation involving multiple documents with non-isomorphic structure (and including, therefore, manuscripts of the Anglo-Saxon penitentials as described above) it is, in most cases, necessary to develop a custom parser. We achieved this by using a locally-developed framework known as BetterXML, which provides an object-oriented Document Object Model (or DOM) for mapping XML data to a meaningful class model. In addition, the framework supports roundtrip code generation, where the code being generated is the XHTML interface itself. To support *interactivity*, JavaScript is used for the presentation layer with the aim of providing a responsive implementation of the interface described above. We believe the lessons learned in developing custom parsers and a human-usable web interface (i.e. going beyond simple HTML web pages) made possible by so-called Web 2.0 technologies will be of value to the broader community exploring digital humanities and is of paramount importance to other research topics, including search and retrieval.

We will demonstrate the operation of the database and identify features under development, including a search tool and the limited use of JPEGs that show the actual condition and layout of the manuscripts so that users can compare them to the facsimile editions.